International Journal of Engineering & Technology, 7 (4.39) (2018) 30-36

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Comparative Analysis of Optimizer Methods for Machine Learning Algorithm using Search Keyword Ad Data

Hyunhee Park[1*], Minjun Ji[2]

[1]Korean Bible University, Department of Computer Software, Nowon-gu Seoul, South Korea, 01757
[2]e-Glue Communications, Researcher, Seocho-gu Seoul, South Korea, 06651
**Corresponding author Email: parkhyunhee@gmail.com*

Abstract

A searching for specific keywords through an online web searching platform (e.g., Google or Naver) is one of the most popular search mechanisms. Therefore, an advertisement on the online web searching platform has become one of the representative advertising marketing mechanisms. In order to advertise a specific keyword, it is necessary to pay the web-search-word to the online web searching platform and participate in the bid system. However, since the ads of online web searching platform are operated privately, it is quite difficult to know the bidding price of specific keywords. In this paper, we compare analysis results of machine learning algorithms with various optimizers to find the targeted rank of specific keywords and the desired ranking on the online web searching platform by using the machine learning algorithms. Particularly, it is quite important to find an appropriate optimization mechanism for the machine learning algorithm because it can derive different results of the applied machine learning algorithm according to the optimization mechanism. Therefore, we propose an appropriate machine learning algorithm with various optimizers by analyzing the web-search-word advertisement data. The ANN of deep learning and regression (i.e., linear, logistic, softmax regressions) algorithms are applied for the machine learning algorithms. In addition, we applied the optimizer mechanisms of Adam, Adagrad, Gradient Descent, Momentum and RMSProp to these algorithms. Extensive simulation results show that the Adam and Adagrad optimizer mechanisms have high test accuracy rate. Specifically, it can be seen that each optimizer mechanism shows quite difference in accuracy rate according to learning rate. Finally, it is necessary to analyze the machine learning algorithm applying various optimizer mechanisms to present the bidding price prediction model of the web-search-word advertisement. In this paper, it makes it possible to predict the optimal bidding price for the web-search-word advertisement.

Keywords: *Web-search-word, Keyword advertisement, Machine learning algorithm, Optimizer mechanism, ANN, Regression*

1. Introduction

Most of a web-search-word advertising data is consumed in the web searching plat such as an online portal. In order to launch an advertisement execution, a real time bidding (RTB) system is increasingly introduced to an advertising auction solution [1, 2]. When an advertiser wants to start launched advertisements through the web searching platform, normally the RTB system will use the web searching platforms such as the Google AdWords and Microsoft Bing Ads. Nowadays, the web advertising platform using KaKao is widely used in South Korea. When an advertiser needs to launch a web-search-word advertisement through the Naver or Kakao web searching platform, the advertising auction is generally executed through the RTB system named the Kakao Ad or Naver Search AD.
The specific web-search-word chosen by the advertisers to represent characteristics of their products or services are associated with an advertisement [3]. It appears as a sponsored link at the result web page in a web searching platform in response to the query of users. Considering the higher the ranking at the result page in a web searching platform, the higher the price should be paid. Because high-ranking search results ultimately leads to high click-through rates. To our knowledge, few researches discuss web-search-word data modeling, hence greatly motivating this research.
In the web-search-word advertising, a return to investment (ROI) is quite popular ratio metric between the net profit and cost of investment resulting from an investment of some resources. In [4], authors investigate that advertisers are maximizing their ROI value across multiple web-search-words in sponsored search auctions, however others bidding with second prices may not select for them. A cost per click (CPC), is an internet advertising model used to direct traffic to websites, in which an advertiser pays a publisher when the ad is clicked. On the other hand, a pay-per-click (PPC) is commonly associated with first-tier search engines. With search engines, advertisers typically bid on keyword phrases relevant to their target market. The weighted-sum method is mainly used. Therefore, the targeted ranking of search keyword is mainly selected by the maximum bidding price, the click rate, and other related factors. In [4], authors provide the periodic patterns in in various statistics including impressions, clicks, bids, and conversion rates such as post-view rate and post-click rate. Therefore, when the user clicks the related link in the web searching results, the CPC-based keyword advertisement has expenses. In the web-search-word advertising, the top ranked keyword is sold by a high bidding price at the top of web searching platform. However, the actual consumption cost of the advertising such as the CPC value can be different from the bidding price offered in the web searching bidding solution. The same offering bid prices can deal with different positions at web

result page due to different advertising models. The bidding price determines not only the price per click, but also the position of the ad in the sponsored search results, consequently costs, revenue and finally profitability of sponsored search. In terms of prediction robustness, the semi-logarithmic model is reported to be much better than others, but it is insufficient to predict ROI due to the problem of its simplicity. In [5],[6], they suggest that under certain conditions, different channels may need to be used to optimize returns to advertising for advertisers and service providers. In addition, authors developed a Bayesian click-through rate prediction model to predict the click-through rate (CTR) of paid search ads. But the generalized linear model with a cumulative Gaussian function associated with ad impressions may exclude several critical input variables such as position.

When a specific search-word advertisement is implemented, several related links may be executed at least 20 for one web-search-word. In addition, the ranking of the web-search-word can be affected by various input factors such as the bidding price and date. This is an example of the bidding price and related ranking: \$10 for 1st rank of the web searching result, \$8 for 2nd rank of the web searching result, \$6 for 3rd rank of the web searching result, and etc. In this case, since the RTB system is performed as a blind competition, it is not possible to know directly the bidding price for the targeted rank. Therefore, it is also difficult to predict bidding price and expected rank in the online web searching platform because bidding price and expected rank can form different time zones.

In this paper, we introduce a feature selection algorithm that executes to find the targeted rank and optimized machine learning algorithms for the prediction model of bidding price. Particularly, we expiscate the results of optimizers for the targeted rank using the regression algorithms and neural networks in the scikit-learn and TensorFlow. Particularly, in order to predict high quality result of prediction appropriate optimization algorithms are applied in the machine learning algorithm. Because it can derive different results of the applied machine learning algorithm according to the optimization mechanism. Therefore, in this paper, we present and analyze the results of applying the various optimization mechanisms by analyzing the web-search-word advertisement data. The ANN of deep learning and regression (i.e., linear, logistic, softmax regressions) algorithms are applied for the machine learning algorithms. In addition, we applied the optimization mechanisms of Adam, Adagrad, Gradient Descent, Momentum and RMSProp to these algorithms. Extensive simulation results show that the Adam and Adagrad optimization mechanisms have high test accuracy rate. Specifically, it can be seen that each optimization mechanism shows quite difference in accuracy rate according to learning rate. Finally, it is necessary to analyze the machine learning algorithm applying various optimization mechanisms to present the bidding price prediction model of the web-search-word advertisement.

2. System Model and Prediction Model for Bidding Price of Web-search-word Data

In order to collect web-search-word data, we use a web crawler kind of spiderbot from client APs that copies web pages for processing by a search engine which indexes the downloaded pages so users can search more efficiently generated web searching results. In this research, when multiple client APs supply the same web-search-word data, they will be automatically saved as one data. Even though the web-search-word data is the same, a price can be different according to time. In particular, web-search-word data varies in bidding price depending on working date such as weekdays and weekends. In addition, the same data has different tendencies according to the daily time. In Fig. 1, the bidding price of the web-search-word data changes over a weekday. On the other hand, the bidding price of the web-search-word data changes over weekends in Fig. 2. It is shown that there is a quite difference between weekday and weekend data depending on the bidding price of specific keyword. As a result of comparing Fig. 1 and Fig. 2, it can be seen that the bidding price of web-search-word data is higher at the weekend. Also, the time value is an important input for the prediction model. In order to apply the time value in the prediction algorithm, we divide one hours into 30 minutes as 48 units into time sets.

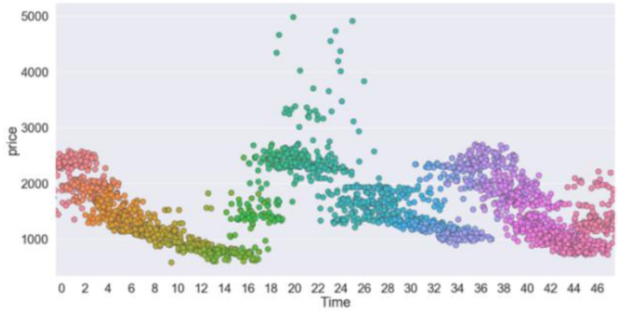

Figure 1: Results of specific web-search-word on weekdays.

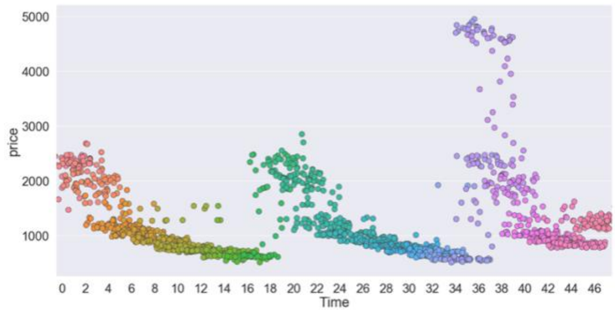

Figure 2: Result of specific web-search-word on weekends.

In the conventional human-based search advertising system, when a user desires to register a specific web-search-word in a targeted rank of the linked web page, the bidding price of the specific web-search-word must be manually calculated. As a result, there is a problem that the system is constructed to set the bidding price based on the simple statistical analysis or manually in order to decide the bidding price of the web-search-word, which requires low accuracy rate and a lot of manpower. In addition, the "keyword bidding method based on the click ranking" which can know the current ranking can confirm the targeted ranking by the update period of the click ranking. However, the update period of the click ranking should depend on the results provided by the web searching platform. For example, the update period of click ranking in the Naver web searching platform is specified to be about 4 hours. As a result, it is difficult to guarantee the ranking of web-search-words to be targeted. Because the web-search-word advertisement is rapidly updated in 15 ~ 60 seconds. In order to solve these problems, we propose a prediction model of targeted rank based on a machine learning algorithm that analyzes the web-search-word data and rankings. In particular, this paper applies various optimizer mechanisms and analyzes the results to improve the accuracy rate of the test results in the machine learning algorithms.

Basically, various optimization mechanisms are applied particularly the gradient descent mechanism [7]. In this paper, we use the five optimization algorithms: Adam, Adagrad, Gradient Descent, Momentum, and RMSProp. In addition, the adaptation of learning rate is the important process to minimize the cost function such as the overshooting and local minimum problems. In order to these problems, finding of the proper learning rate with the various optimization mechanisms is quite important. In this paper, we apply the input data to 151: 2 for price, 48 for time set, 1 for weekdays or weekends, 100 for one keyword name with word2vec. In order to apply the multiple layered algorithm such as the deep learning algorithm, the hidden dimension value is considered 256 and 512 classes.

3. Simulation Results and Discussion

A web-search-word data is composed with a database using MySQL based relational database management system (RDBMS). First, since the bidding price result of the web-search-word depends on the time, the web-search-word is analyzed according to the time. To do this, the 24 hours is composed to the units of 30 minutes to construct a web-search-word data set based on 48 data sets. The time data sets are applied to the one-hot-encoders. In addition, since the Korean language of web-search-word data needs to be processed, the web-search-word data is converted into numerical vector values using the word embedding. The input set as x contains the value of the bidding price and time, and the output data set as y determines the targeted rank value of the web-search-word data. The web-search-word bidding data set consists of training and test data set as 70% and 30%, respectively.

3.1. Results of ANN Algorithms for the Prediction Model

As a result of deriving the training results for the ANN algorithm, it can be seen that when the Adam optimizer mechanism is applied, the value fluctuates little and the accuracy rate increases step by step as the epoch increases. However, it can be shown that an appropriate learning rate is applied to the Adam optimizer mechanism to find the optimum result. When the Adam optimizer mechanism is applied to the ANN algorithm, the accuracy rate of the training data is the highest when the learning rate is 0.001. In particular, when the learning rate is 0.01, the accuracy rate of training data is reached quickly, but when the learning rate is 0.001, a higher accuracy rate of training data is obtained. In addition, when the Adam optimizer mechanism is applied to the ANN algorithm, the value of the cost function is deduced. As the Epoch increases, the cost value becomes minimum. As the learning rate of the training data is 0.001, it shows the highest accuracy rate in the Adam optimizer mechanism and the cost function result reaches the minimum value when the learning rate is 0.001 in the Adam optimizer. However, it can be seen that the learning rate converges to the minimum more quickly when the learning rate is 0.01 in the Adam optimizer mechanism. Figure 3 compares the accuracy rate of ANN algorithms by applying various optimizer mechanisms to test data sets. Similar to the training data set described above, the Adam optimizer mechanism showed the highest accuracy rate when the learning rate was applied to 0.01 compared to other optimizer mechanisms. Figure 4 shows the result of the cost function according to the various optimizer mechanisms. When the learning rate of the Adam optimizer mechanism is 0.01, it converges to the minimum value as same as the accuracy rate of the test data set as shown in Fig. 4.

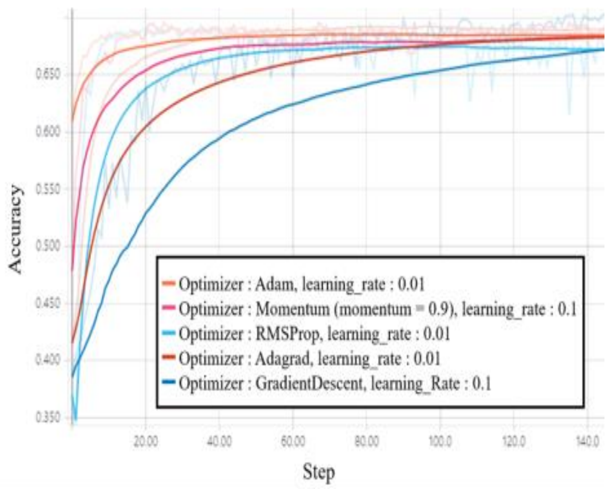

Figure 3: Test accuracy rate of ANN algorithm with various optimizer mechanisms and different learning rates.

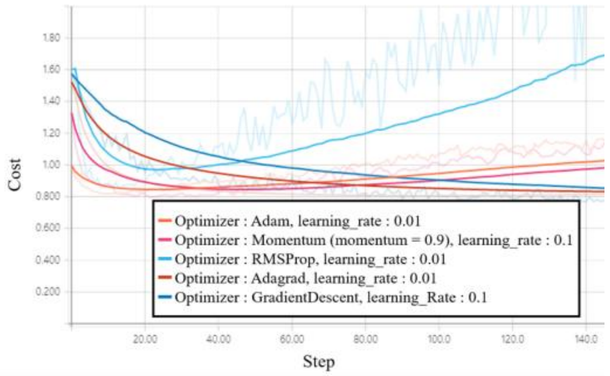

Figure 4: Test cost function of ANN algorithm with various optimizer mechanisms and different learning rates.

Table 1: Results of the ANN algorithm

ANN	Learning rate	Epoch	Evaluation cost	Evaluation accuracy rate
Adagrad	0.1	46	0.792119	69.36%
	0.01	52	0.821699	68.86%
	0.001	150	0.957203	64.88%
	0.0001	150	1.511379	42.08%
Adam	0.1	16	1.687831	32.21%
	0.01	10	0.818790	68.64%
	0.001	10	0.817781	68.46%
	0.0001	81	0.822337	68.51%
Gradient Descent	0.1	100	0.806881	69.16%
	0.01	120	0.968841	64.59%
	0.001	150	1.503623	41.75%
	0.0001	150	1.752510	35.42%
Momentum	0.1	101	0.993491	68.54%
	0.01	73	0.811343	68.96%
	0.001	150	0.927609	65.86%
	0.0001	150	1.504111	41.74%
RMSProp	0.1	5	1.685417	32.64%
	0.01	150	3.077648	65.49%
	0.001	150	1.815077	66.50%
	0.0001	150	0.795354	69.58%

Table 1 shows more specific results for the ANN algorithm. It can be seen that accuracy rate and cost function results vary depending on the learning rate. Particularly, when the ANN algorithm is applied, the accuracy rate of the training data and the test data are the highest when the learning rate is 0.01 with the Adam optimizer mechanism. Particularly, the test cost reaches the minimum value when the learning rate is 0.01 and epochs are 400 in the Adam optimizer mechanism. In this case, as the epochs increase, the cost value increases again. The results of the Adagrad mechanism are quite high comparing with the Adam optimizer mechanism. In particular, the Adagrad optimizer mechanism shows the highest accuracy rate when the learning rate is 0.1. However, in comparison with the Adam and Adagradoptimizer mechanisms, the Adam optimizer mechanism is more dominant because the epoch needs to at least 6,200 for high test accuracy rate.

3.2. Results of Linear Regression Algorithms for the Prediction Model

As a result of deriving the accuracy rate of training data for the linear regression algorithm, it can be seen that when the Adam optimizer mechanism is applied, the accuracy rate of the training data is high. It can be seen that an appropriate learning rate is applied to the Adam optimizer mechanism to find the optimum result. When the Adam optimizer mechanism is applied to the linear regression algorithm, the accuracy rate of the training data is the highest when the learning rate is 0.1. In addition, when the learning rate is 0.1 of the Adam optimizer mechanism, the accuracy rate of training data is reached quickly. Furthermore, when the Adam optimizer mechanism is applied to the linear regression algorithm, the value of the cost function is deduced. As the Epoch increases, the cost value becomes minimum. Figure 5 compares the accuracy rate of linear regression algorithms by applying various optimizer mechanisms to test data sets. Similar to the training data set described above, the Adam optimizer mechanism showed the highest accuracy rate when the learning rate was applied to 0.1 compared to other optimizer mechanisms. Figure 6 shows the result of the cost function according to the various optimizer mechanisms. When the learning rate of the Adam optimizer mechanism is 0.1, it converges to the minimum value as same as the accuracy rate of the test data set as shown in Fig. 6. Particularly, the Adagrad optimizer mechanism, which showed the high accuracy rate in the ANN algorithm, shows very low accuracy rate in the linear regression algorithm. In addition, the gradient decent optimizer mechanism has the lowest accuracy rate and the momentum optimizer mechanism predicts the accuracy rate very slowly.

Table 2 shows more specific results for the linear regression algorithm. As mentioned above, it can be seen that accuracy rate and cost results vary depending on the learning rate. Particularly, when the linear regression algorithm is applied, the accuracy rate of the training data and the test data are the highest when the learning rate is 0.1 with the Adam optimizer mechanism. Particularly, the test cost function reaches the minimum value when the learning rate is 0.1 and epochs are 400 in the Adam optimizer mechanism. In particular, the Adagrad optimizer mechanism shows the highest accuracy rate when the learning rate is 0.1. However, in comparison with the Adam and Adagrad optimizer mechanisms, the Adam optimizer mechanism is more dominant because the epoch needs to at least 6,200 for high test accuracy rate.

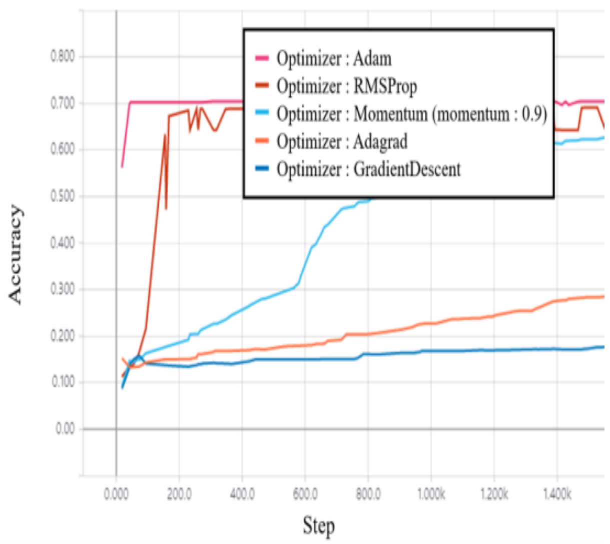

Figure 5: Test accuracy rate of linear regression algorithm with various optimizer mechanisms and different learning rates.

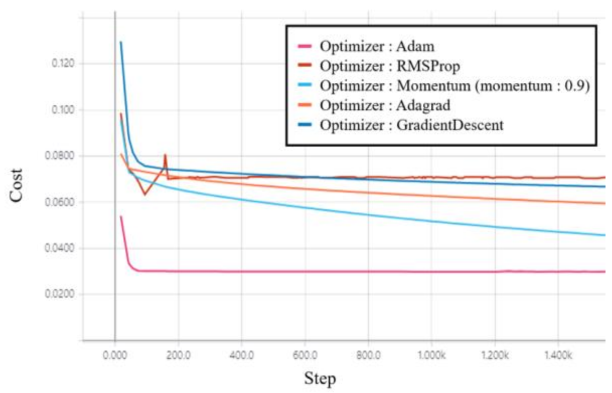

Figure 6: Test cost function of linear regression algorithm with various optimizer mechanisms and different learning rates.

Table 2: Results of the linear regression algorithm

Linear	Learning rate	Epoch	Test cost function	Test accuracy rate
Adagrad	0.1	2600	0.030341	70.26%
	0.1	6200	0.029972	70.42%
	0.01	11300	0.035177	69.84%
	0.001	15000	0.060441	25.68%
	0.0001	15000	0.075148	13.59%
Adam	0.1	400	0.029946	70.42%
	0.1	15000	0.033306	70.01%
	0.01	400	0.029945	70.42%
	0.001	1700	0.029966	70.42%
	0.0001	3200	0.035165	70.42%
Momentum	0.1	600	0.031157	70.42%
	0.01	8100	0.030359	70.26%
	0.001	15000	0.046125	62.30%
	0.0001	15000	0.066890	17.23%
RMSProp	0.1	200	4.104390	30.90%
	0.1	5000	4.099177	30.74%
	0.01	300	0.071020	64.29%
	0.01	5700	0.070834	64.62%
	0.001	1500	0.030387	70.26%
	0.0001	2600	0.035867	70.42%

3.3. Results of Logistic Regression Algorithms for the Prediction Model

As a result of deriving the accuracy rate of training data for the logistic regression algorithm, it can be seen that when the Adam optimizer mechanism is applied, the accuracy rate of the training data is quite high. On the other hand, it can be shown that an appropriate learning rate is performed to the Adam optimizer mechanism to find the optimal performance. When the Adam optimizer mechanism is applied to the linear regression algorithm, the accuracy rate of the training data is the highest when the learning rate is 0.1.

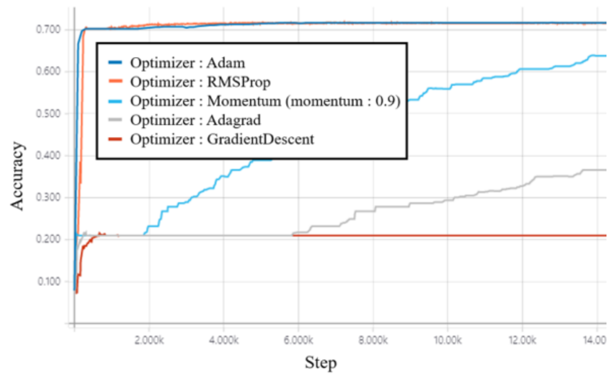

Figure 7: Test accuracy rate of logistic regression algorithm with various optimizer mechanisms and different learning rates.

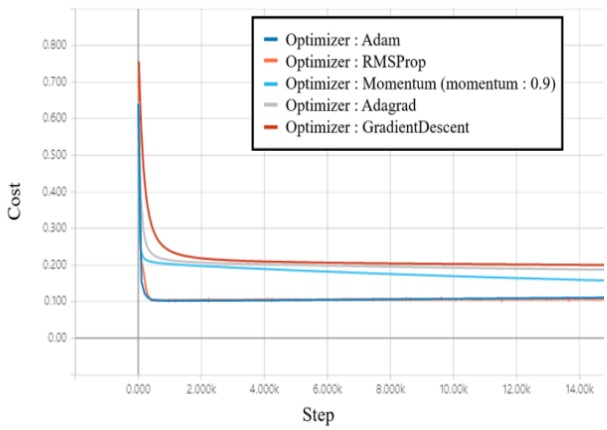

Figure 8: Test cost function of logistic regression algorithm with various optimizer mechanisms and different learning rates.

In addition, when the learning rate is 0.1 of the Adam optimizer mechanism, the accuracy rate of training data is reached quickly. Furthermore, when the Adam optimizer mechanism is applied to the linear regression algorithm, the value of the cost function is deduced. As the Epoch increases, the cost value becomes minimum. In addition, the logistic regression algorithm also shows a fairly high accuracy rate with the RMSProp optimizer mechanism. This shows that the Adam and RMSProp optimizer mechanisms have similar low-cost values in the test cost function results. Figure 7 compares the accuracy rate of logistic regression algorithms by applying various optimizer mechanisms to test data sets. Similar to the training data set described above, the Adam optimizer and RMSProb mechanisms showed the highest accuracy rate when the learning rate is applied to 0.1 compared to other optimizer mechanisms. Figure 8 shows the result of the cost function according to the various optimizer mechanisms. When the learning rate of the Adam optimizer and RMSProb optimizer mechanisms are 0.1, it converges to the minimum value as shown in Fig. 8. Particularly, the Adagrad optimizer mechanism, which showed the high accuracy rate in the ANN algorithm, shows very low accuracy rate in the logistic regression algorithm. In addition, the gradient decent optimizer mechanism has the lowest accuracy rate and the momentum optimizer mechanism predicts the accuracy rate very slowly.

Table 3: Results of the logistic regression algorithm

Logistic	Learning rate	Epoch	Test cost	Test accuracy rate
Adagrad	0.1	15000	0.121743	70.67%
	0.01	15000	0.186684	37.37%
	0.001	6800	0.234853	20.96%
	0.0001	6100	0.626258	7.54%
Adam	0.1	2000	0.105446	71.67%
	0.01	5500	0.104454	71.67%
	0.001	2200	0.117127	70.26%
	0.001	12800	0.104426	70.75
	0.0001	11100	0.136360	69.51%
Gradient Descent	0.1	13830	0.159745	63.88%
	0.01	900	0.243878	20.96%
	0.001	8200	0.249209	20.96%
	0.0001	15000	0.484747	12.34%
Momentum (0.9)	0.1	5700	0.117454	70.26%
	0.01	15000	0.157152	63.63%
	0.001	900	0.243352	20.96%
	0.0001	8200	0.249150	20.96%
RMSProp	0.1	5990	0.172766	63.79%
	0.01	4200	0.105285	71.67%
	0.001	12000	0.103113	71.25%
	0.0001	12300	0.127838	69.59%

Table 3 shows more specific results for the logistic regression algorithm. It can be shown that the accuracy rate and cost results vary depending on the various learning rate. Specifically, when the logistic regression algorithm is applied, the accuracy rate of the training data and the test data are the highest when the learning rate is 0.1 with the Adam optimizer and RMSProb optimizer mechanisms.

Particularly, the test cost function reaches the minimum value when the learning rate is 0.1 and epochs are 2,000 in the Adam optimizer mechanism. In addition, the results of the Adagrad mechanism are only high accuracy rate when the learning rate is 0.1. In the other cases (i,e., 0.01, 0.001, and 0.0001 of learning rates), the accuracy rate is quite low. Furthermore, the logistic regression algorithm shows the similar accuracy rate as the Gradient Descent optimizer and Momentum optimizer mechanisms with the linear regression algorithm.

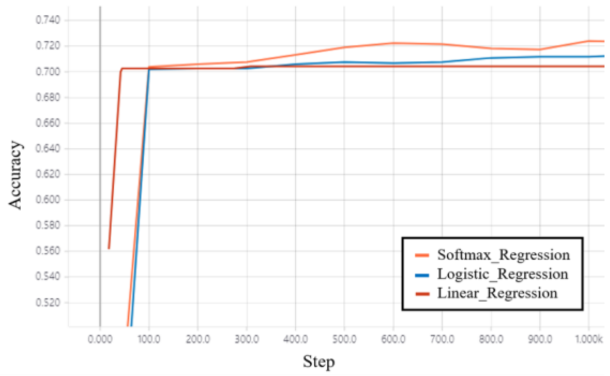

Figure 9: Comparison of test accuracy rate results for the regression algorithms with best optimizer mechanisms and learning rates.

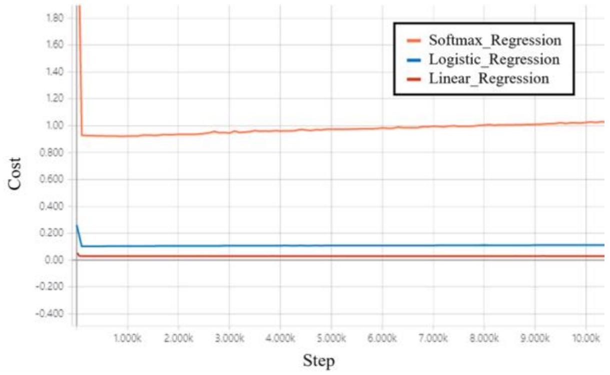

Figure 10: Comparison of test cost function results for the regression algorithms with best optimizer mechanisms and learning rates.

In principal, the logistic regression is used for one class separation. In case of a softmax regression algorithm is implemented as a multinomial classification. Therefore, the cost value is large compared to the linear regression and the logistic regression algorithms. Nonetheless, we compared the three algorithms because the accuracy rate of the softmax regression algorithm is rapidly converging among three algorithms. Figure 9 shows that the softmax regression algorithm reaches the highest accuracy rate. The optimizer applied to the softmax regression algorithm is the Adam optimizer mechanism and the learning rate is 0.1. The optimizer with the high accuracy rate of the linear regression algorithm is the Adam optimizer mechanism and the learning rate is 0.01. The optimizer applied to the logistic regression algorithm is the Adam optimizer mechanism and the learning rate is 0.1. In the case of the logistic regression algorithm, the Adam and RMSProb optimizer mechanisms show almost similar results, but because the Adam optimizer mechanism shows more superiority in the cost value, we choose the Adam optimizer mechanism for the logistic regression algorithm. As shown in Fig. 10, the softmax algorithm is the most disadvantage in the cost value. This is because, as mentioned above, the softmax algorithm is implemented for the multinomial classification.

3.4. Results of Softmax Regression Algorithms for the Prediction Model

As a result of deriving the accuracy rate of training data for the softmax regression algorithm, it can be seen that when the Adam optimizer mechanism is applied, the accuracy rate of the training data is high. In this case, it can be shown that an appropriate learning rate is applied to the Adam optimizer mechanism to find the optimal performance. In addition, when the learning rate is 0.1 of the Adam optimizer mechanism, the accuracy rate of training data is reached quickly. On the other hand, when the Adam optimizer mechanism is applied to the softmax regression algorithm, the value of the cost function is deduced. As the Epoch increases, the cost value becomes minimum. In addition, the softmax regression algorithm also shows a fairly high accuracy rate withthe RMSProp optimizer mechanism. This shows that the Adam and RMSProp optimizer mechanisms have similar low-cost values in the test cost function results.

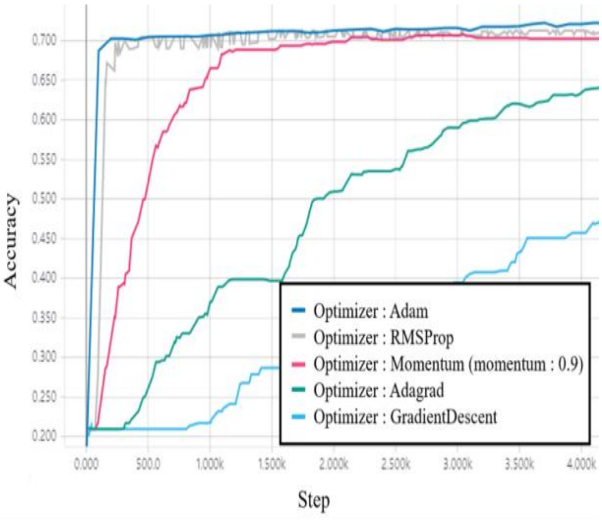

Figure 11: Test accuracy rate of softmax regression algorithm with various optimizer mechanisms and different learning rates.

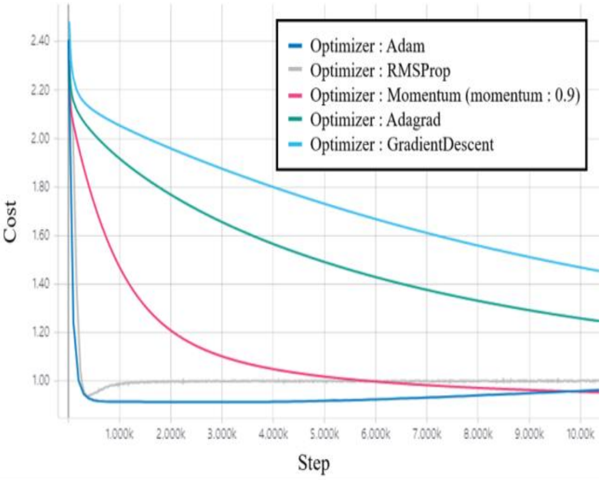

Figure 12: Test cost function of softmax regression algorithm with various optimizer mechanisms and different learning rates.

Figure 11 compares the accuracy rate of softmax regression algorithms by applying various optimizer mechanisms to test data sets. Similar to the training data set described above, the Adam optimizer and RMSProb mechanisms showed the highest accuracy rate when the learning rate was applied to 0.1 compared to other optimizer mechanisms. However, the RMSProb optimizer mechanism shows very frequent fluctuations in the Fig. 11. Figure 12 shows the result of the cost function according to the various optimizer mechanisms. When the learning rate of the Adam optimizer and RMSProb optimizer mechanisms are 0.1, it converges to the minimum value as same as the accuracy rate of the test data set as shown in Fig. 12. Particularly, the Adagrad optimizer mechanism, which showed the high accuracy rate in the ANN algorithm, shows very low accuracy rate in the logistic regression algorithm. In addition, the gradient decent optimizer mechanism has the lowest accuracy rate and converges very slowly.

Table 4 shows more specific results for the softmax regression algorithm. It can be shown that accuracy rate and cost function results vary depending on the learning rate. Particularly, when the softmax regression algorithm is applied, the accuracy rate of the training data and the test data are the highest when the learning rate is 0.1 with the Adam optimizer mechanism.

Table 4: Results of the softmax regression algorithm

Softmax	Learning rate	Epoch	Test cost	Test accuracy rate
Adagrad	0.1	15000	0.909653	70.75%
	0.01	15000	1.142566	70.34%
	0.001	15000	1.963206	32.06%
	0.0001	15000	2.276862	21.62%
Adam	0.1	4800	0.969772	72.91%
	0.01	5000	0.918634	72.33%
	0.001	15000	0.965082	71.67%
	0.0001	15000	1.024288	70.59%
Gradient Descent	0.1	15000	0.935226	70.34%
	0.01	12100	1.391469	68.85%
	0.001	15000	2.004422	28.67%
	0.0001	5100	2.336776	20.96%
Momentum	0.1	15000	0.896408	71.67%
	0.01	15000	0.935221	70.34%
	0.001	12000	1.394947	68.85%
	0.0001	15000	2.004485	28.67%
RMSProp	0.1	15000	2.756302	59.40%
	0.01	11300	1.003106	71.50%
	0.001	11700	0.969739	71.25%
	0.0001	15000	0.998624	70.34%

3.5. Proposed One-Gap Error Range

As a result, the average accuracy rate of the training data is estimated to be 73%, and the average accuracy rate of the test data is estimated to be 70.86%. This is because the lack of data sets caused the underfitting problem. The rank of web-search-word advertising data has various range of ranking value from first to fifteenth. Therefore, it is difficult to predict the exact ranking value even though the appropriate machine learning algorithm is applied. In this paper, we define a one-gap error range for a ranking accuracy rate. It means if the current ranking of the web-search-word data is predicted as the third rank, it is possible to allow the one-gap error range such as second and fourth ranks as the correct prediction results. Because the web-search-word data in the online page has change quite quickly the rank value in real time. When the one-gap error range is allowed, the accuracy rate of the prediction is greatly improved. The accuracy rate of the ANN algorithm is 92%, the accuracy rate of the linear regression algorithms is 93%, the accuracy rate of the logistic regression algorithm and the softmax regression algorithm are 94% and 96%, respectively. As a result, it can be seen that accuracy rate is improved when the one-gap of the error is set for the ranking prediction.

4. Conclusion and Future Work

We perform the targeted rank prediction models of web-search-word data by various machine learning algorithms and optimizer mechanisms. As the deep neural network, the result of ANN model is showed as 77% accuracy rate of the training data and 69% accuracy rate of the test data, respectively. In the case of the regression algorithms, the result of linear regression algorithm is showed as 74% accuracy rate of the training data and is 71% accuracy rate of the test data, and the result of logistic regression algorithm is showed as 75% accuracy rate of the training data and is 72% accuracy rate of the test data, and the result of softmax algorithm is showed as 76% accuracy rate of the training data and is 73% accuracy rate of the test data on average depending on the various optimizer algorithms, respectively. In order to solve the underfitting problem, we propose the one-gap error range for the ranking accuracy rate. After that, the accuracy rate of predicted ranking can be increased. In order to solve this underfitting problem, we gather the web-search-word data more sufficiently and compare the results of these prediction models.

Acknowledgment

This work was supported by the Research Grant from Korean Bible University (2017).

References

[1] Hou, L. (2015). A Hierarchical Bayesian Network-Based Approach to Keyword Auction, IEEE Trans. on Engineering Management, 62(2), 217–225.
[2] Lauritzen, S. (1995) The EM algorithm for graphical association models with missing data, Comput. Statistics Data Anal., 19(2), 191–201.
[3] Jansen, B.J., Zhang, M., & Schultz, C. D. (2009). Brand and its effect on user perception of search engine performance, Journal of the association for information science and technology, 60(8), 1572–1595.
[4] Auerbach, J., Galenson, J., & Sundararajan, M. (2008). An empirical analysis of return on investment maximization in sponsored search auctions, Las Vegas, NV, USA: International Workshop Data Mining Audience Intell. Ad.
[5] Jerath, K., Ma, L., Park, Y., & Srinivasan, K. (2011). A Position Paradox in sponsored search auctions, Marketing Sci., 30(4), 612–627.
[6] Kingma,D.&Ba,J. (2008). ADAM: A Method for Stochastic Optimization,San Diego, CA, USA: International Conference on Learning Representations.
[7] Duchi, J., Hazan, E., & Singer Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization, The Journal of Machine Learning Research, 12(2011), 2121–2159.